PB86-154879

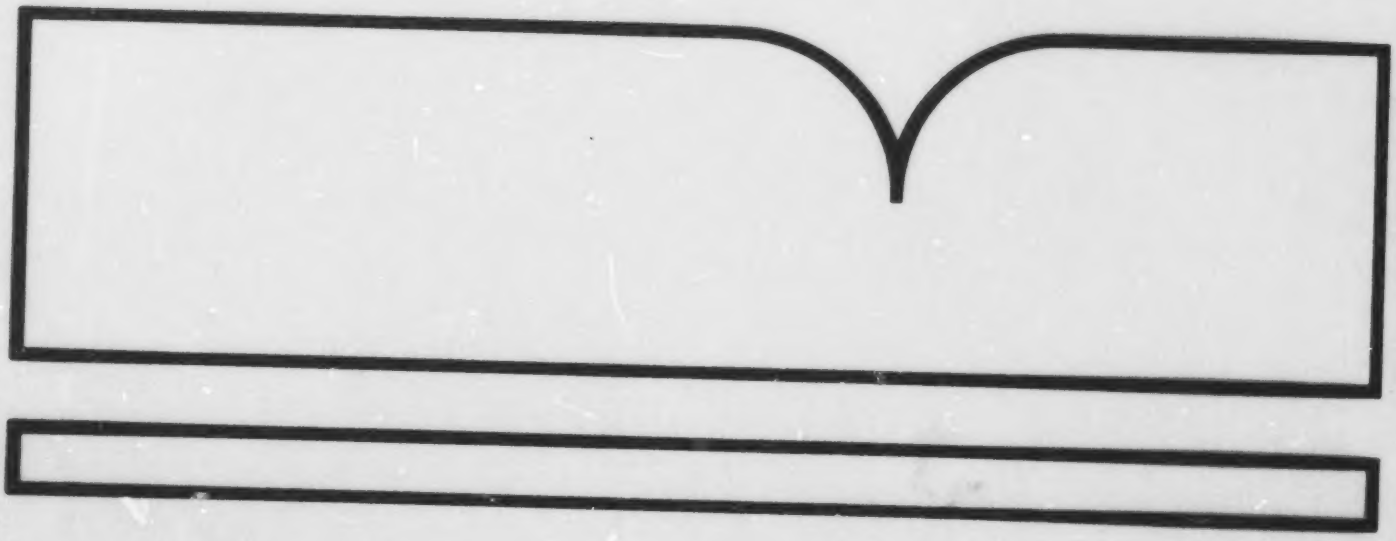
Weilbull Function as a Dose-Response Model to Describe Ozone Effects on Crop Yields

North Carolina State Univ. at Raleigh

Prepared for

Corvallis Environmental Research Lab., OR

1985



U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA/600/J-85/260	2.
3. RECI PB86-154879	
4. TITLE AND SUBTITLE The Weibull Function as a Dose-Response Model to Describe Ozone Effects on Crop Yields (Journal Version)	5. REPORT DATE 1985
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Rawlings, J.O. and W.W. Cure	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS North Carolina State University P.O. Box 5186 Raleigh, NC 27650	10. PROGRAM ELEMENT NO.
	11. CONTRACT/GRANT NO. IAG AD-12-F-1-490-2
12. SPONSORING AGENCY NAME AND ADDRESS Environmental Research Laboratory Office of Research and Development U.S.Environmental Protection Agency Corvallis, OR 97333	13. TYPE OF REPORT AND PERIOD COVERED Journal Article
	14. SPONSORING AGENCY CODE EPA 600/02

15. SUPPLEMENTARY NOTES
1985. Crop Science 25:807-814, September-October. Project Officer: E.M.Preston FTS 420-4636

16. ABSTRACT

Field experiments have demonstrated that ozone (O_3) pollution can reduce yields of important crop species. The assessment of the impact of O_3 is facilitated by the use of a family of dose response models that covers the range of responses observed in field O_3 dose response studies. The Weibull dose response model is a flexible, biologically realistic, model that has been found useful for this purpose. The purposes of this study are i) to discuss analyses of dose response data using the Weibull model and its extensions, II) to identify some of the problems inherent in the use of a nonlinear dose response model, and iii) to show application of the model to two sets of experimental data.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

18. DISTRIBUTION STATEMENT Release to Public	19. SECURITY CLASS *(This Report)* Unclassified	21. NO. OF PAGES 9
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE

EPA Form 2220-1 (9-73)

Reprinted from *Crop Science*
Vol. 25, Sept.-Oct. 1985, p. 807-814

EPA/600/J-85/260
JOURNAL ARTICLE

The Weibull Function as a Dose-Response Model to Describe Ozone Effects on Crop Yields[1]

J. O. Rawlings and W. W. Cure[2]

ABSTRACT

Field experiments have demonstrated that ozone (O_3) pollution can reduce yields of important crop species. The assessment of the impact of O_3 is facilitated by the use of a family of dose response models that covers the range of responses observed in field O_3 dose response studies. The Weibull dose response model is a flexible, biologically realistic, model that has been found useful for this purpose. The purposes of this study are i) to discuss analyses of dose response data using the Weibull model and its extensions, ii) to identify some of the problems inherent in the use of a nonlinear dose response model, and iii) to show application of the model to two sets of experimental data. The basic Weibull model is a modification of the Weibull probability distribution function. Extensions of the basic model allow for class variables and other continuous variables in the experimental design. A bivariate form of the Weibull allows for independent effects of two pollutants. The use of the Weibull model with class variables (cultivars) is demonstrated using a 1977 winter wheat (*Triticum aestivum* L.) experiment grown at Raleigh, NC involving four cultivars exposed to four levels of O_3 pollution. Tests of homogeneity over cultivars of the proportional yield responses show differential cultivar response to O_3. The bivariate Weibull model is demonstrated using a 1981 soybean [*Glycine max* (L.) Merr.] experiment involving one cultivar exposed to the factorial set of treatments consisting of six levels of O_3 and four levels of SO_2 pollution. The test of lack of fit of the bivariate Weibull model was nonsignificant, supporting the hypothesis that O_3 and SO_2 were acting independently. The percent predicted relative yield losses and their standard errors are given for two of the wheat cultivars and for the soybean cultivar.

Additional index words: **Response curves, Nonlinear models, Regression, Homogeneity of regressions, Ozone pollution, *Triticum aestivum* L., *Glycine max* (L.) Merr.**

DATA from field experiments indicate ozone (O_3) can reduce yields of important crop species, and it is estimated that O_3 is responsible for up to 90% of the crop losses associated with air pollutants in the USA (4,5,6,7). Models relating O_3 concentrations to crop yields are required to quantify the impact of present O_3 concentrations and to assess benefits of alternative O_3 standards.

Open-top field chambers provide a means of exposing field grown crops to different gaseous pollutants for extended periods of time without extensive disruption of the crop microenvironment (3). As part of the National Crop Loss Assessment Network (NCLAN), experiments using these chambers have begun to generate O_3 dose-yield response data for a variety of crops growing in different parts of the USA (6,7). The first attempts to use these data employed simple linear equations to model the response to O_3 (7). Equations for individual cultivars were averaged to approximate the behavior of species regardless of degree of similarity of responses.

Although linear equations were adequate for cultivars of many crops, definite nonlinear yield responses to O_3 were indicated in data for cultivars of such important crops as corn (*Zea mays* L.) and wheat (*Triticum aestivum* L.) (7). Several alternative models that would accommodate the curvilinear responses observed were investigated: higher degree polynomials, plateau-linear, logistic, and S-shaped response curves based on normal and Weibull probability distribution functions. Many other possibilities for models exist, including empirical transformations of variables that linearize the response (1). Higher-degree polynomials, quadratic in particular, tended to give false peaks of yield performance at low doses in cases where the major impact of O_3 was at higher doses. Therefore, results overestimated the yield loss due to O_3 since the false peak usually occurred near ambient level of O_3.

Except for this apparent deficiency of the quadratic model, any of several three parameter models were not distinguishable in terms of goodness of fit due to the limited amount of information in any one data set. It was desirable, however, to have a common model that could be used to characterize all dose response data sets. The model based on the Weibull probability function has proved useful in many situations, including modeling crop disease progression (8). The Weibull model was chosen to represent the O_3 dose-crop yield response relationships for a variety of reasons:

1. It has a flexible form that covers the range of responses observed.
2. The form of the Weibull is biologically realistic.
3. The parameters of the Weibull have straightforward interpretations.
4. The Weibull provides direct estimates of proportional yields.
5. Tests of homogeneity of proportional yield responses over data sets are readily accomplished.

The purpose of this paper is to present the Weibull model as modified for the study of O_3 effects, to discuss some enhancements to the basic model, and to demonstrate its use.

THE WEIBULL MODEL

The cumulative probability distribution function for the Weibull is given by

$$F(x;\mu,\sigma,c) = 1 - \exp\{-[(x-\mu)/\sigma]^c\}, \qquad [1]$$

where $\mu < x < \infty$, $\sigma > 0$, and $c > 0$. This function increases monotonically from zero (when x equals its lower bound, μ) to 1 (when x equals infinity). Different biological modes of action for O_3 could be argued to rationalize use of the Weibull function, but our interest stems from its convenient mathematical form for proportional yield dose response models of O_3 damage. For this purpose, we have chosen to express the response model in terms of the proportion of the

[1] Paper no. 9279 of the journal series of the North Carolina Agric. Res. Service, Raleigh, NC 27695. Received 18 June 1984.
[2] Professor, Dep. of Statistics, and research associate, Dep. of Botany, North Carolina State Univ., Raleigh, NC 27695, respectively.

NTIS is authorized to reproduce and sell this report. Permission for further reproduction must be obtained from the copyright owner.

yield remaining at dose x, $H(x)=1-F(x;\mu,\sigma,c)$, and we have assumed that $\mu=0$. To provide for actual rather than proportional yield levels, a parameter α is introduced to represent the hypothetical maximum yield at zero O_3 dose, giving

$$Y_i = \alpha \exp[-(x_i/\sigma)^c] + \epsilon_i \qquad [2]$$
$$= \alpha H(x_i) + \epsilon_i$$

as the basic Weibull yield response model, where Y_i is yield and x_i is an appropriate measure of O_3 exposure. The measure of O_3 exposure used herein is the 7 h day^{-1} seasonal mean of the O_3 concentrations, $\mu L\ L^{-1}$ (ppm). (Exposures are for 7 h [1000 to 1700 h] each day for most of the growing season for the specific crop.) The exponential part of the model, labeled as $H(x)$, characterizes the proportional yield response. The parameter σ has a direct interpretation as the dose at which yield is reduced to 0.37α. The parameter c is a dimensionless parameter controlling the rate at which damage is incurred, i.e., the shape of the response curve. If $c=1$, the Weibull function becomes the familiar exponential curve with the relative rate of yield loss being $1/\sigma$ for all x. With increasing c, the impact of the pollutant is increasingly concentrated near $x=\sigma$, with the relative rate of loss at that point being c/σ. Thus, larger values of c give a distinct plateau effect to the response curve.

The Weibull is nonlinear in the parameters. The nonlinear model fitting routine, PROC NLIN in SAS,[3] has been used to estimate the parameters α, σ, and c. Implicit in this decision to fit the nonlinear model directly is the assumption of additive, independent errors with common variance. It is reasonable to expect variances of the residuals to decrease with decreasing yield, in which case some modification of the analysis would be in order. Analyses to date have not indicated heterogeneous variances within an experiment to be a series problem. A decision was made to avoid, if possible, transformations of the dependent variable, either to stabilize variances or to linearize response, to retain the original units of measure. No simple transformation would linearize the model unless some parameters were forced to take specific values, such as $c=1$ or $\alpha=1$. Requiring $c=1$ destroys the flexibility of the Weibull model, and $\alpha=1$ is unrealistic if yield is to be the variable analyzed. Rescaling yield by dividing, for example, by the largest observed yield is equivalent to forcing the largest yield to be the estimate of α. This would give positive bias to estimates of yield at $x=0$ and would overestimate the impact of O_3.

The independence of errors assumption is not satisfied when data sets from different experiments are being combined. Observations within a data set tend to be positively correlated by random environmental differences between experiments. This creates no problem if the treatments, including the cultivars used, are the same over all experiments. However, if treatments differ over experiments, such as experiments being combined to compare responses of different cultivars, there is no way to separate fixed cultivar by O_3 interaction effects from random variation in response caused by random elements of site differences. This is a general problem of combining experiments from different environments and is not unique to NCLAN experiments or to the use of the Weibull model. Similarly, data from subplot treatments within a split-plot experiment will tend to be positively correlated. This correlation is properly handled by the conventional split-plot analysis of variance, but it causes confusion in choice of error terms for tests of sums of squares obtained from fitting nonlinear models. A full analysis would take into account the correlated error structure by using a generalized least squares estimation. Most nonlinear estimation programs, however, are not designed to handle general weighting.

We have ignored the correlated error structure except when the conventional split-plot analysis was used for data from a single experiment. This caused some loss in precision in estimates of the parameters (but should not introduce any biases) and, as mentioned, caused some confusion in selection of appropriate error terms for testing H_0. In combining data over cultivars, or levels of any treatment factor, where the cultivars have been tested in different sites, one must recognize that tests of significance are only guidelines since interactions of fixed effects (cultivars) can never be separated from interactions of the random environmental differences. The latter are properly part of experimental error for making inferences to the geographical region.

EXTENSIONS OF THE WEIBULL MODEL

The basic Weibull model can be extended in many ways to incorporate other effects in experiments. The procedure we have adopted is to expand the α parameter by adding terms to represent other effects as one does in linear additive models. Here, in contrast to purely additive models, the exponential part of the Weibull, $H(x)$, is multiplied by the extended α-term. The interpretation of this form of the model is that the sum of all other effects determines the level of performance in "clean" air, and the impact of the O_3 is represented by a proportional yield loss determined by concentration. The following are the more common expansions of the Weibull model we have used.

Inclusion of Class Variables

Three different class variables have been considered. Some of the NCLAN experiments have used nonchambered plots (exposure to ambient O_3) to obtain a measure of the effect of the exposure chamber. The effect of the chamber has been included in the Weibull model by appropriate definition of a dummy variable to reflect the presence or absence of the chamber. Similarly, block effects, cultivar effects, or both have been included by appropriate use of dummy variables.

The chamber effect is handled with a single dummy variable defined to take the value of one when the plot has a chamber and zero when there is no chamber. The model becomes

$$Y_i = (\alpha_1 + \alpha_2{}^*CH_i)\, H(x_i) + \epsilon_i, \qquad [3]$$

[3] SAS is the registered trademark of SAS Institute Inc., Cary, NC.

where CH_i is the chamber dummy variable. Parameter α_1 is the yield at zero dose in the absence of a chamber, and α_2 is the change in yield at zero dose due to presence of the chamber. Thus, $(\alpha_1 + \alpha_2)$ is the yield at zero dose for the chambered plots. The significance of the chamber effect, α_2, can be evaluated with a t test. It is important to note that the chamber effect in the model causes the dose response function to be estimated only from the chambered plots since all the nonchambered plots will have very nearly the same seasonal average exposure to O_3.

For class variables with more than two levels, such as cultivar or block effects, one dummy variable must be defined for each level of the class variable. Thus, the dummy variable defined for a cultivar, for example, takes the value one if the observation is from that cultivar and the value zero otherwise. There must be a cultivar parameter corresponding to each dummy variable. Thus, the model becomes

$$Y_i = (\alpha_1 Z_1 + \alpha_2 Z_2 + \ldots + \alpha_b Z_b)\, H(x_i) + \epsilon_i \quad [4]$$

where Z_j is the dummy variable for the jth cultivar and α_j is the corresponding parameter. Thus, α_1 is the expected yield at zero O_3 for cultivar 1, etc. This extended model with cultivar dummy variables is used for testing homogeneity over cultivars of the proportional yield responses to O_3. The model allows for cultivar differences in yield but requires all cultivars to have the same proportional yield loss when exposed to the same level of O_3.

One further extension of the Weibull model allows each cultivar to have its own response to O_3, including its own value of σ and c. This model could be written, using cultivar dummy variables Z_j defined above, as

$$Y_i = Z_1\, \alpha_1\, H_1(x) + Z_2\, \alpha_2\, H_2(x) + \ldots + \epsilon_i, \quad [5]$$

where each $H(x)$ function has a σ and c parameter unique for that cultivar. Fitting this model is more easily accomplished by fitting the simple Weibull model separately to each cultivar data set. Comparison of the sums of squares for Eq. [4] and Eq. [5] provides the basis for testing homogeneity of proportional yield loss functions, the $H(x)$, over cultivars.

Inclusion of a Continuous Variable

We have considered two kinds of continuous variables. Several NCLAN experiments have used yields from companion plots as covariates to account for local environmental variation. Other covariates may be of value in other experiments. Other NCLAN experiments have included levels of a second pollutant as a treatment factor. The α parameter can be extended to accomodate the covariate or terms of a polynomial response curve in those cases where it is appropriate for the second pollutant. The model becomes

$$Y_i = (\alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots) H(x_i) + \epsilon_i, \quad [6]$$

where the X_i are values of the covariates or powers of the dose variable for terms of the polynomial.

The Bivariate Weibull

Should the response to a second pollutant also have a Weibull form, a bivariate Weibull model having an $H(x)$ function for each pollutant can be used:

$$Y_i = \alpha\, H1(x1_i)\, H2(x2_i) + \epsilon_i. \quad [7]$$

The $H1$ term and $x1$ are the Weibull loss function and dose, respectively, for one pollutant; $H2$ and $x2$ are the corresponding quantities for the other pollutant. This model assumes the two pollutants act independently in the sense that each increment of pollutant acts on the undamaged yield "units" according to its own damage function with no interplay from the other pollutant. Any significant interplay between pollutants, either synergistically or antagonistically, would be reflected in a significant lack of fit for the model. It should be noted that this is a multiplicative model. Independence of two factors in a multiplicative model is not the same as absence of interaction in the more conventional additive model. Thus, significant interaction in the additive model is not inconsistent with independence in the multiplicative model and vice-versa.

METHODS OF ANALYSIS

The following outline of the method of analysis presumes that each set of data comes from a balanced experiment in which each of the O_3 treatments is equally replicated in an appropriate experimental design. It is assumed that one of the treatments is an ambient air treatment without a chamber and that only data from chambered plots are to be used in estimating the proportional yield response function, $H(x)$. In the event the experiments are not balanced, the methods presented here must be modified to account for the imbalance.

Analysis of Variance

The appropriate analysis of variance was run using all data to obtain estimates of experimental errors. These were used (i) in estimation of standard errors of parameter estimates, (ii) in (approximate) tests of significance of the Weibull model results, and (iii) in weighting individual experiments, if necessary, when combining experiments. The partitions of the treatment sum of squares using only chambered data were also needed and were obtained from a second analysis of variance.

A few NCLAN experiments consisted of a single replicate of a two-factor factorial, for example, O_3 by sulfur dioxide (SO_2). In these cases, experimental error was estimated by the residual mean square after fitting a bivariate polynomial response surface to the data. The degree of the polynomial response surface was determined by various methods of variable selection in multiple regression such as Mallows C_p statistic (2).

Fitting the Weibull Model in a Single Factor Experiment

In this section we consider a single data set where O_3 was the only treatment factor. Programs for fitting nonlinear models are iterative and require "starting values" for the parameters. In most cases, the derivatives of the model with respect to each of the parameters are also required. We found that the fitting procedure was fairly insensitive to choice of starting values for the Weibull model as long as they were reasonable and the data set was reasonably well behaved, i.e., it showed a basic Weibull form. The starting value for the parameter σ was easily read from a plot of the treatment means against O_3 dose as the (approximate) dose at which the yield had been reduced to one-third that

at zero dose; this sometimes required extrapolation of the curve. The starting value for the parameter c was set to about 1 or 3, depending on whether the yield response was essentially linear or showed a plateau type response. The starting value for α was set near the highest yield. Convergence from these starting values was no problem as long as there was sufficient response to the pollutant to define the Weibull parameters.

Some nonlinear fitting programs have a derivative free method. It has been our impression that the results from the nonlinear estimation have been somewhat better using derivatives. The derivatives for the basic Weibull model written in SAS code are

$$\text{DER}\cdot A = H(x) \quad [8]$$

$$\text{DER}\cdot s = A^*H(x)^*(c/s)^*[(x/s)^{**}c] \quad [9]$$

$$\text{DER}\cdot c = -A^*H(x)^*[(x/s)^{**}c]^*\log(x/s)\,, \quad [10]$$

where A and s are used for α and σ. For the extended models, the DER·A is replaced with derivatives with respect to each of the parameters defined in the extended α term. Each of these derivatives has the form $H(x)$*(dummy variable for the parameter). The A in the other derivatives is replaced with the extended α term.

Block effects need not be included in the Weibull model when the data are balanced and, unless there is specific interest in estimating the chamber effect, it is easier to fit the models using only data from the chambered plots. In the following, we assume the data are balanced and that only chambered data are used; $SS(O_3)$ will refer to the treatment sum of squares from the analysis of variance of these data. To make the Weibull model sum of squares, SS(MODEL), fit into the analysis of variance, we subtract the sum of squares attributable to an overall mean (CF). Then, $SS(\text{Model}|\mu) = SS(\text{MODEL}) - CF$ is a component of the $SS(O_3)$. The lack of fit (lof) of the simple Weibull model is measured by $SS(\text{lof}) = SS(O_3) - SS(\text{Model}|\mu)$, which has t-3 degrees of freedom (t is the number of O_3 treatment levels). $SS(\text{Model}|\mu)$ has two degrees of freedom. Thus, both significance of the Weibull model and lack of fit of the Weibull model can be tested (approximately) using the appropriate error term estimated in the analysis of variance.

The standard errors of parameter estimates provided by PROC NLIN, for example, are computed using the residual mean square from the nonlinear analysis as the estimate of experimental error. This will be biased by any omitted terms in the Weibull model such as the block effects. This can be corrected by multiplying standard errors of parameter by the square root of the ratio of the analysis of variance error mean square, MSE, to the residual mean square from the nonlinear analysis, MSRES.

Fitting the Weibull Model in a Two Factor Experiment

In this section we consider the analysis of a single experiment but where the treatment set consists of a factorial set of treatments composed of O_3 levels and a class variable that we assume to be cultivars. We are interested in using the Weibull model to describe total variation in the O_3 by cultivar two-way table of data.

It is of interest to minimize the number of distinct Weibull functions needed to characterize crop yield response to O_3 by combining cultivars into groups that respond similarly. For this purpose, it is the proportional yield response that is important, not the actual yield level. Therefore, we want to combine cultivars according to similarity of the fitted $H(x)$ functions.

There are three nested Weibull models that can be used to partition the total sum of squares from the two-way O_3 by cultivar table, $SS(O_3 + \text{Cult})$. They are in order of increasing generality:

1. The simple Weibull requires all cultivars to have the same response to O_3 and the same yield level. Thus, there are three parameters, α, σ, and c. The model sum of squares is denoted by SS(simple) and has three degrees of freedom.
2. The common Weibull requires all cultivars to have the same proportional yield response, common σ and c, but allows different yield levels (different α). Therefore, this model has $b+2$ parameters, where b is the number of cultivars. The sum of squares is denoted by SS(common) and has $b+2$ degrees of freedom.
3. The full Weibull allows every cultivar to have its own response curve including different values of σ and c. Thus, there are $3b$ parameters, three for each of the b cultivars. The model sum of squares is denoted by SS(full) and has $3b$ degrees of freedom.

The partitioning of the total sum of squares for the O_3 by cultivar table is as follows:

1. $SS(\text{simple}|\mu) = SS(\text{simple})—CF$ has two degrees of freedom and measures the contribution of the simple Weibull model to $SS(O_3)$.
2. $SS(\text{common}|\text{simple}) = SS(\text{common})—SS(\text{simple})$ has b-1 degrees of freedom and measures the improvement in the model from allowing each cultivar to have its own α_i.
3. $SS(\text{full}|\text{common}) = SS(\text{full})—SS(\text{common})$ has $2(b\text{-}1)$ degrees of freedom and measures the heterogeneity of the proportional dose response curves over the b cultivars, i.e., the heterogeneity of the σ and the c.
4. $SS(\text{lof}) = SS(O_3+\text{Cult})—SS(\text{full})$ has $b(t\text{-}3)$ degrees of freedom and measures the lack of fit of the full Weibull, the model allowing each cultivar to have its own response.

The above partitioning is based on a multiplicative model, and although the partitions appear to be similar to those in the analysis of variance, they do not match. This difference in partitioning creates confusion in determining the appropriate errors for testing if the data set is from a split-plot experiment. The $SS(\text{simple}|\mu)$ is part of $SS(O_3)$; thus, the appropriate error for testing would be the error appropriate for testing O_3 effects (Error a in the NCLAN studies where O_3 levels are whole plot treatments). $SS(\text{common}|\text{simple})$ and $SS(\text{full}|\text{common})$ are partitions of $SS(\text{Cultivar}) + SS(\text{Cult}^*O_3)$. The former is measuring cultivar differences but on a multiplicative scale and clearly is not equal to or part of the cultivar sum of squares. The latter is primarily measuring interaction. When cultivars are the subplot treatment, as is the case with the NCLAN experiments, the appropriate error for testing both would be the subplot error (Error b). If the cultivars were whole plot treatments, however, the appropriate errors would not be clear. SS(lof) as given above involves parts of both the whole plot and the subplot analyses. However, it is possible to partition out the sum of squares due to the lack of fit of the simple Weibull to the O_3 treatment means. The remainder of SS(lof) is to be tested with Error b of the split-plot analysis.

Since different yield levels of the cultivars are expected and since we are primarily interested in the proportional

yield losses, there may be no need to separate *SS*(simple) from *SS*(common). The primary questions to be answered with this partitioning are (i) is there significant lack of fit of the Weibull model, in which case other models need to be considered; and (ii) can the response of the cultivars to O_3 be adequately represented by a single $H(x)$ response function.

Fitting a Common Weibull Model Over Data Sets

Two situations need to be distinguished in combining data sets. The first is where an experiment has been repeated in several random environments with the objective of estimating a common Weibull response curve for the set of environments. The second situation is where the experiments at the different sites differ by a second treatment factor, such as different cultivars, and the objective is to assess homogeneity of responses to O_3 over this other factor. The problems occur in the latter when the second factor is considered to be a fixed effect.

The first case differs from the clasical combining of experiments over random environments in that the levels of the O_3 treatments will never be the same in different sites since the O_3 is being added to ambient levels. If the O_3 levels over environments differ by more than trivial amounts, the conventional partitioning of the treatment by environment two-way table is not appropriate. The apparent site effects and site by O_3 interaction effects will be caused, in part, by the different levels of O_3 and, hence, the site by O_3 interaction mean square will not be a "clean" estimate of the appropriate error as in the usual "random environments" case. For this situation, we suggest that the partitioning of the two-way table presented above be abbreviated to include fitting only the common Weibull model, a common σ and c but an α_i for each environment, and the residual mean square from this two-way table be used as the estimate of error appropriate for the combined Weibull. That is, *SS*(res) = *SS*(O_3+Envir) − *SS*(commonμ) and will have (bt-b-2) degrees of freedom.

The second case where sites may include differences in a second treatment factor that is not a random effect, cultivars for example, will frequently also present the problem of unequal numbers of O_3 treatments. In this case, the Weibull model partitioning of the *SS*(O_3+Envir) could proceed as described above for the *SS*(O_3+Cult) in the single experiment. The total degrees of freedom will be 1 less than the sum of the number of O_3 treatments in each site if the numbers of treatments are not equal. The testing discussed above does not apply to these comparisons, and, as discussed elsewhere, any tests of significance must be interpreted with caution. Any observed heterogeneity cannot a priori be ascribed solely to the cultivar differences because of confounding with other environmental differences. Likewise, any test of lack of fit is a priori expected to be larger than the pooled error because of the random environmental differences between the sites. If it is assumed that there is no true lack of fit of the full Weibull model, the *MS*(lof) would be a reasonable error for testing if it has sufficient degrees of freedom.

A final comment on combining data sets. In some cases, the data sets to be combined have heterogeneous error variances; even the units on which yield is based may differ over experiments. In these cases, we have chosen to standardize data from each data set by dividing each observation by the square root of the error mean square from the analysis of variance for that experiment. This, in effect, gives equal weight to each observation. It does not affect the interpretation of the σ and c. We have continued to use the mean squares computed from the weighted data as discussed in the above analyses of the unweighted data.

Predicted Relative Yield Loss (PRYL)

The objective of developing the yield dose response models is to provide an estimate of the yield losses expected from various O_3 levels. Again, it is the proportional yield loss that is of interest, and this information is provided by $H(x)$. Let Xo be the base level of O_3 from which yield losses are to be measured, and let Xr be the postulated new level of O_3. Then, PRYL = $1 - H(Xr)/H(Xo)$ where the $H(\cdot)$ functions are evaluated using the estimates of the parameters σ and c.

Variances

The approximate asymptotic variance for the estimated proportional yield at dose x, $H(x)$, is given by

$$\text{Var}[\text{H}(x)] = [H(x)]^2\{[A(x)]^2 \text{ var}(\hat{c}) + [B(x)]^2 \text{ var}(\hat{\sigma}) - 2\,A(x)\,B(x)\,\text{cov}(\hat{c},\hat{\sigma})\}, \quad [10]$$

where $\hat{c}$ and $\hat{\sigma}$ are the parameter estimates, $A(x)$ equals $(x/\hat{\sigma})^{\hat{c}}\,ln(x/\hat{\sigma})$, and $B(x)$ equals $(\hat{c}/\hat{\sigma})(x/\hat{\sigma})^{\hat{c}}$.

The approximate variance of the predicted relative yield loss, PRYL, is given by

$$\text{Var(PRYL)} = [H(Xr)/H(Xo)]^2\{[A(Xr)-A(Xo)]^2 \text{ var}(\hat{c}) + [B(Xr)-B(Xo)]^2 \text{ var}(\hat{\sigma}) - 2[A(Xr)-A(Xo)][B(Xr)-B(Xo)]\,\text{cov}(\hat{c},\hat{\sigma})\}. \quad [11]$$

USE OF THE WEIBULL

Analysis of two published experiments, one with winter wheat and one with soybeans [*Glycine max* (L.) Merr.], demonstrate the use of the Weibull model. The 1977 winter wheat experiment was designed to characterize the response of four cultivars of winter wheat to O_3 (5). Cultivars Blueboy II (BY), Coker 47-27 (CO), Holly (HO), and Oasis (OS) were grown in a split-plot design with cultivars as the subplot treatments. Whole plot treatments were arranged in a randomized complete block (RCB) design with four blocks and five O_3 treatments, four in open top chambers and one nonchambered ambient air plot (AA). Charcoal filtering (CF) was used to reduce O_3 below ambient levels, and O_3 was added to nonfiltered air (NF) to create higher levels. The four chambered O_3 treatments and the ambient air treatment gave seasonal 7 h day^{-1} mean O_3 levels of 0.03 (CF), 0.06 (NF), 0.10, 0.13, and 0.06 (AA) μL L^{-1}, respectively.

The 1981 soybean experiment was designed to characterize the yield response of the cultivar Davis to two pollutants, O_3 and sulfur dioxide (SO_2) (4). There were two complete replicates of 25 treatments: all combinations of six levels of O_3 and four levels of SO_2 in open-top chambers plus one nonchambered plot (AA). The six O_3 treatments gave seasonal 7 h day^{-1} mean O_3 levels of 0.025 (CF), 0.055 (NF), 0.069, 0.086, 0.106, and 0.125 μL L^{-1}. The four SO_2 treatments gave seasonal 4 h day^{-1} mean SO_2 levels of 0.0, 0.026, 0.085 and 0.367 μL L^{-1}. The AA plot had seasonal means of 0.053 μL L^{-1} for O_3 and 0.0 μL L^{-1} for SO_2.

Table 1. Randomized complete block analysis of variance for the wheat data, the Weibull partitioning of the O_3 sum of squares, and the Weibull parameter estimates for the wheat data.

Source	df	Cultivar: Blueboy	Coker	Holly	Oasis
		Mean squares			
		Analyses of variance—additive model			
Blocks	3	0.9118	0.4971	0.9902	0.3801
O_3	3	2.8404	3.5649	2.3756	1.1999
Error†	12	0.1969	0.1186	0.1768	0.1639
		Analysis of variance—Weibull model			
Model (μ)‡	2	4.2556	5.3138	3.5634	1.7951
lof§	1	0.1000	0.0672	0.0000	0.0095
		Weibull parameter estimates¶			
$\hat{a}$		5.884	5.194	4.951	4.484
		(0.223)	(0.286)	(0.167)	(0.204)
$\hat{\sigma}$		0.175	0.171	0.156	0.186
		(0.023)	(0.016)	(0.013)	(0.040)
$\hat{c}$		3.22	2.06	4.95	3.20
		(1.33)	(0.68)	(2.03)	(1.86)

† Error is estimated using all data including nonchambered plots. Blocks and O_3 are estimated only from chambered plots.
‡ *SS* (Model μ) = *SS* (Weibull model) − CF using only chambered data. CF is the sum of squares due to correcting for the mean.
§ *SS* (lof) = *SS* (O_3) − *SS* (Model μ) measures lack of fit.
¶ Standard errors of estimates are shown in parentheses. These are asymtotic standard errors computed from the nonlinear program and adjusted to reflect the analysis of variance error mean square.

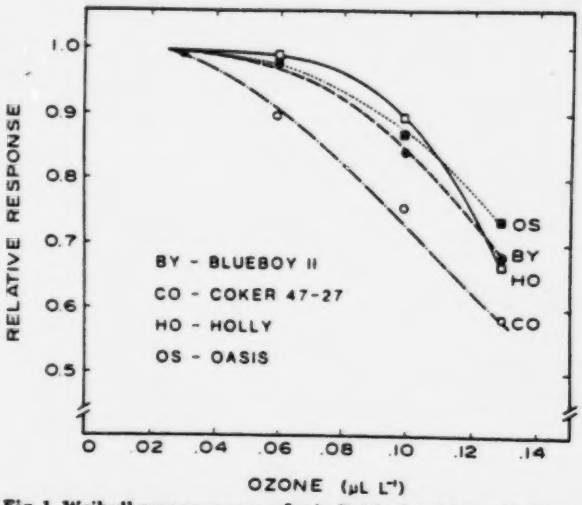


Fig. 1. Weibull response curves for individual cultivars of wheat. The common Weibull response curve is slightly below the curve for BY. All yields are scaled to show 100% yield at $x = 0.025$ μL L^{-1}. The O_3 concentration is the seasonal 7 h day^{-1} (1000–1700 h) mean.

Analysis of Wheat Experiment

The wheat data are used first to illustrate the analysis of a single factor experiment by regarding each cultivar as a separate experiment (Table 1). The RCB analysis of variance for each cultivar shows a clear indication of significant O_3 effects. The fitting of the simple Weibull model accounts for almost all the O_3 effects. (The tests for lack of fit in this example have very low power because of the limited number of levels of the O_3 treatment.) The proportional yield function for each cultivar is given by the $H(x)$ function using the estimates of σ and c. Thus, for the cultivar Coker,

$$H(x) = \exp(-(x/0.171)^{2.06}, \quad [12]$$

The predicted relative yield (PRY) using a reference level of O_3 of 0.025 μL L^{-1} is obtained by the ratio of $H(x)$ to $H(0.025)$. Thus, for the cultivar Holly, $H(0.025)$ equals 0.9811 and

$$\text{PRY} = H(x)/0.9811, \quad [13]$$

and rescales Holly to 100% yield at x equals 0.025 μL L^{-1}. The predicted relative yields are shown in Fig. 1 for the four cultivars. The predicted relative

Table 2. Percent predicted relative yield losses, PRYL, from O_3 (standard errors in parentheses) for the wheat cultivars Holly and Coker and for the soybean cultivar Davis. Predictions are based on the fitted Weibull models and are relative to the yield at 0.025 μL L^{-1}.

O_3 conc†	PRYL: Wheat: Holly	Wheat: Coker	Soybeans: Davis
μL L^{-1}	%		
0.03	0.02 (0.05)‡	0.9 (0.7)	4.0 (0.9)
0.04	0.1 (0.2)	3.1 (2.0)	11.3 (2.1)
0.05	0.3 (0.7)	5.9 (3.1)	17.9 (2.7)
0.06	0.9 (1.4)	9.2 (4.0)	23.9 (2.8)
0.07	1.9 (2.3)	13.0 (4.6)	29.4 (2.7)
0.08	3.6 (3.5)	17.3 (4.8)	34.5 (2.6)
0.09	6.3 (4.5)	21.9 (4.7)	39.1 (2.4)
0.10	10.4 (5.5)	26.8 (4.4)	43.4 (2.4)
0.11	16.3 (5.7)	31.8 (3.9)	47.3 (2.4)
0.12	24.1 (4.9)	37.0 (3.5)	51.0 (2.6)
0.13	33.4 (4.7)	42.2 (3.6)	54.3 (2.9)

† Ozone concentrations are seasonal 7 h day^{-1} means in μL L^{-1}
‡ Standard errors are computed using equation (11). Var ($\hat{\sigma}$) and Var ($\hat{c}$) are squares of standard errors in Table 1. Cov ($\hat{c}$, $\hat{\sigma}$) are −0.02357, −0.008857, and 0.00391 for Holly, Coker, and Davis, respectively, and are computed from the correlation matrix in the nonlinear program results.

Table 3. Split plot analysis of variance of wheat data and Weibull model partitioning of variance in ozone × cultivar 2-way table using weighted results for each cultivar.†

Source	df	*SS*	MS
	Analysis of variance – additive model		
Blocks	3	45.3423	15.1141
O_3	3	180.6292	60.2097
Error *a*‡	12	34.8184	2.9015
Cultivars	3	42.2187	14.0729
Cult × O_3	9	15.0951	1.6772
Error *b*‡	45	18.5452	0.4121
	Weibull model partitioning of (O_3 + Cult)		
Total (O_3 + Cult)	15	237.9430	
simple μ	2	180.5238	90.2619
common | simple	3	47.2411	15.7470
full | common	6	8.4916	1.4153
lack of fit	4	1.6865	0.4216
	Common Weibull model parameter estimates§		
$\hat{\sigma}$ = 0.172 (0.018)			
$\hat{c}$ = 3.09 (1.05)			

† Weighting is done by dividing all observations from each cultivar by the square root of the error mean square for that cultivar from Table 1. This would normally not be done with a self-contained experiment but is presented here to illustrate the weighting for combining different experiments.
‡ Error *a* and Error *b* are taken from the split plot analysis of all data including nonchambered plots to capitalize on the additional degrees of freedom.
§ Standard errors (in parentheses) from the nonlinear program have been adjusted to Error *a* (above) by multiplication by the square root of the ratio Ea/(NLIN residual mean square).

Table 4. Analysis of variance and estimates of the parameters for the bivariate Weibull model for the soybean O_3 by SO_2 study.

Source	df	SS	MS
Analysis of variance—additive model			
Blocks	1	382	382
O_3	5	220 681	44 136
SO_2	3	76 948	25 649
$O_3 \times SO_2$	15	17 781	1 185
Exp. error	23	24 489	1 065
Analysis of $O_3 \times SO_2$ with bivariate Weibull model			
Total ($O_3 + SO_2$)	23	315 411	
(Weibull |μ)†	4	301 971	75 493
Residual	19	13 440	707

Bivariate Weibull parameter estimates‡

$\hat{\alpha} = 549$ (80)
$\hat{\sigma}_1 = 0.129$ (0.017)
$\hat{c}_1 = 0.914$ (0.284)
$\hat{\sigma}_2 = 0.848$ (0.230)
$\hat{c}_2 = 1.072$ (0.343)

† SS (Weibull |μ) = SS (MODEL) for the bivariate model minus the correction factor (CF).
‡ Standard errors of the estimates are shown in parentheses.

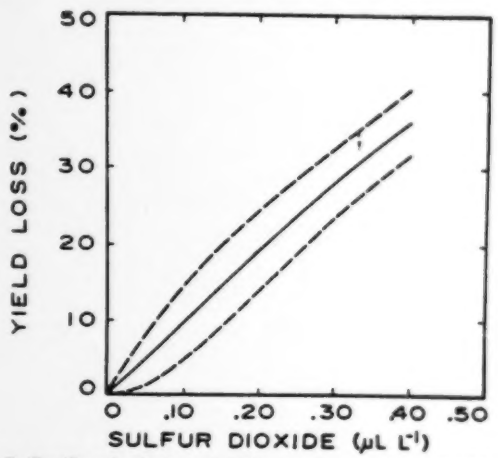


Fig. 2. Predicted relative yield losses from sulfur dioxide for the Davis cultivar of soybeans relative to the yield at zero SO_2 (shown with ± SE). The SO_2 concentration is the 4 h day^{-1} (1000–1400 h) seasonal mean.

yield losses (PRYL) are obtained by subtracting these values from 1. The variances of PRY and PRYL are the same. The PRYL (%) for the most sensitive cultivar, Coker, and most resistant, Holly, are shown with their estimated standard errors in Table 2.

The split-plot analysis of variance of the wheat data (Table 3), was run on the weighted data. Weighting for each cultivar was by the inverse of the square root of the respective error mean squares in Table 1, for example, $1/\sqrt{0.1186}$ for the Coker cultivar. (The weighting here is to illustrate combining experiments with heterogeneous variances and not because of any indication of heterogeneity of errors in Table 1). The analysis of variance shows a cultivar by O_3 interaction on the additive scale, F equals 4.07.

Inspection of the Weibull parameter estimates in Table 1 and Fig. 1 suggests that the cultivar Coker behaves quite differently from the other three cultivars. The Weibull model partitioning of the Total (O_3+Cult) sum of squares (Table 3) gives no indication of a lack of fit of the full Weibull model, which allows each cultivar to have its own response curve, $F = 0.4216/0.4121 = 1.02$. The test of the homogeneity of the σ_j and c_j over cultivars, $F = 1.4153/0.4121 = 3.43$, is significant and is consistent with the interaction test in the analysis of variance; the four cultivars are not showing a homogeneous response to O_3.

Consequently, on the basis of these tests, one would either use a different response curve for each cultivar or investigate further which subsets of the cultivars behave similarly. Alternatively, if a single response curve must be used for the four cultivars, the lack of homogeneity of the responses must be considered as part of the prediction error of O_3 damage. The parameter estimates for the common response curve are shown at the bottom of Table 3.

Analysis of Soybean Experiment

The analyses of variance for the additive model and for the bivariate Weibull model are shown in Table 4. The analysis shows significant O_3 and SO_2 effects but no indication of interaction between the two pollutants on the additive scale. The bivariate Weibull model, with only four parameters, leaves a residual mean square even smaller than the RCB experimental error. With no significant lack of fit of the model, it can be concluded that the two pollutants are acting independently on the multiplicative scale. The impact of each pollutant can be described by its $H(x)$ function irrespective of the level of the other pollutant.

$$H1(x1) = \exp(-(x1/0.129)^{0.914}, \quad [14]$$

for O_3, and

$$H2(x2) = \exp(-(x2/0.848)^{1.072}, \quad [15]$$

for SO_2. The PRYL with their standard errors are shown in Table 2 for O_3 and Fig. 2 and for SO_2.

DISCUSSION

The Weibull model has been used to characterize the relative yield response to O_3 for several NCLAN experiments involving many different crops (6, 7). Only a few data sets have shown significant lack of fit to the model, and in those cases, the lack of fit resulted from one treatment response being very much out of line insofar as showing a Weibull-type response. (Modifications of the Weibull model or alternative models will need to be investigated if these few "non-Weibull" type responses are repeatable in future studies.) Arguments could be presented for a Weibull probability distribution of threshold sensitivities to justify the use of this model, but it is not our intention to suggest that this is the biologically meaningful model. The Weibull model was adopted because of its flexibility and relative ease of use and interpretation. There are many other mathematical functions that from a goodness of fit point of view would be as satisfactory.

The Weibull function starts at α for zero dose and asymptotes to zero, or 100% yield loss, as levels of dose get large. The parameter σ indicates the dose at which 63% of yield has been lost relative to α. For

the wheat cultivars, estimates of this critical dose varied from 0.156 μL L^{-1} to 0.186 μL L^{-1}. The c parameter controls the plateau characteristic of the response curve. With the Holly cultivar, for example, very little loss of yield was realized until the O_3 dose exceeded 0.06 μL L^{-1}. On the other hand, Coker had lost 9% of its yield at that point, relative to the yield at 0.025 μL L^{-1}. The parameter α must be interpreted with some caution when c is small, near 1, and when, as with O_3, it is not possible to have a true zero treatment level. In these cases the estimated α is an extrapolation beyond the data and becomes highly variable with small c. If c is large, the plateau nature of the curve gives little yield change in the extrapolation to zero dose and α is more meaningful.

A final comment is needed on the choice of treatment levels when one anticipates fitting nonlinear models such as the Weibull. It is important that the number and range of dose levels be sufficient to allow dominant features of the response curve to be expressed. In general, however, some precision is lost when an excessive number of distinct levels of the pollutant is used (assuming the same total amount of effort). In the case of the Weibull, preliminary studies on optimum designs for the estimation of PRYL indicate major losses in precision of the estimates if the upper dose levels fall much below σ. This is particularly important for the plateau type responses. For example, only 5% of the yield loss occurs below 0.5σ if $c=4$. If $\sigma=0.18$, as estimated in the wheat data, O_3 dose levels well above realistic levels are needed to make optimum use of resources. The highest O_3 treatment level in the wheat experiment was 0.13 μL L^{-1}; the precision would have been improved with higher doses. In the soybean experiment, the highest level of O_3, 0.125 μL L^{-1}, was very nearly equal to $\hat{\sigma}=0.129$, and $\hat{c}$ was close to 1 so that the response curve was well covered. This is reflected in smaller standard errors (for PRLY) for soybeans than for wheat (Table 2). The highest level of SO_2 in the soybean study was less than half $\hat{\sigma}$, and, again, the standard errors are larger than for soybeans (see standard error bands in Fig. 2).

SUMMARY

The Weibull model has proved to be a useful model for characterizing the effects of O_3 on crop yields. The Weibull function has a flexible form that covers a wide range of monotonically decreasing responses that might be considered realistic for many biologically damaging pollutants. The basic model as used herein has three parameters, all of which are easily interpreted.

ACKNOWLEDGMENTS

We wish to thank Dr. H. S. Heagle for permission to use the results from his soybean and wheat studies in this manuscript. Although the research described in this article has been funded in part by the U.S. Environmental Protection Agency through the Interagency Agreement no. AD-12-F-1-490-2 to the U.S. Department of Agriculture, it has not been subjected to the EPA's required peer and policy review and thus does not necessarily reflect the views of the agency.

REFERENCES

1. Box, G.E.P., and P.W. Tidwell. 1962. Transformation of the independent variables. Technometrics 4:531–550.
2. Draper, Norman, and Harry Smith. 1981. Applied regression analysis. 2nd ed. John Wiley and Sons, New York.
3. Heagle, A.S., D.E. Body, and W.W. Heck. 1973. An open-top field chamber to assess the impact of air pollution on plants. J. Environ. Qual. 2:365–368.
4. ———, W.W. Heck, J.O. Rawlings, and R.B. Philbeck. 1983. Effects of chronic doses of ozone and sulphur dioxide on injury and yield of soybeans in open-top field chambers. Crop Sci. 23:1184–1191.
5. ———, S. Spencer, and M.B. Letchworth. 1979. Yield response of winter wheat to chronic doses of ozone. Can. J. Bot. 57:1999–2005.
6. Heck, W.W., R.M. Adams, W.W. Cure, A.S. Heagle, H.E. Heggestad, R.J. Kohut, L.W. Kress, J.O. Rawlings, and O.C. Taylor. 1983. A reassessment of crop loss from ozone. Environ. Sci. Technol. 17:572A–581A.
7. ———, O.C. Taylor, R. Adams, G. Bingham, J. Miller, E. Preston, and L. Weinstein. 1982. Assessment of crop loss from ozone. J. Air Pollut. Control Assoc. 32:353–361.
8. Pennypacker, S.P., H.D. Knoble, C.E. Antle, and L.V. Madden. 1980. A flexible model for studying plant disease progression. Phytopathology 70:232–235.

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

This article documents research that was conducted under the auspices of the U.S. Environmental Protection Agency (USEPA). As an arm of the government, the USEPA has a paid up, non-exclusive, irrevocable, worldwide license to reproduce, prepare derivative works, and distribute to the public all of the material contained herein, and to grant others acting on its behalf the right to do so. The USEPA grants to the National Technical Information Service, Department of Commerce, the right to act as its agent in the distribution of this material.